PB90-113143

Statistical Issues for Seedling Studies in Air Pollution Research

(U.S.) Corvallis Environmental Research Lab., OR

28 Sep 88

U.S. Department of Commerce
National Technical Information Service
NTIS

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA/600/D-89/110	2.
3. REC PB90-113143	
4. TITLE AND SUBTITLE Statistical issues for seedling studies in air pollution research	5. REPORT DATE
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Charles E. Peterson	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS NSI Technology Services Corp., Corvallis, OR	10. PROGRAM ELEMENT NO.
	11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS US Environmental Protection Agency Environmental Research Laboratory 200 SW 35th Street Corvallis, OR 97333	13. TYPE OF REPORT AND PERIOD COVERED Symposium Paper
	14. SPONSORING AGENCY CODE EPA/600/2

15. SUPPLEMENTARY NOTES IN Proceedings of the IUFRO Air Pollution Conference, Interlaken, Switzerland, 2-8 October 1988, Birmensdorf, pp.496-498

16. ABSTRACT

By the very nature that seedlings represent the future forests, they are an important population for exposure studies. The experimental conditions for seedling studies cover a wide range of treatment combinations, facilities, and exposure regimes. The experimental utilization of experimental studies with tree seedlings requires the appropriate design and statistical concepts employed with any population response study. These issues are reviewed for their relevance to exposure studies and contribution toward studies of forest/tree response to pollutants. In order to realize the potential strengths of these studies and further enhance the intrinsic value of seedling research, the planning of experiments should include both the relevant population of interest and an explicit definition (*a priori*) of all response variables and covariates. The experimental designs should provide for pretreatment measurements, repeated measures within growing seasons, and treatments or measurements which span more than one growing season.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT Release to Public	19. SECURITY CLASS (This Report) Unclassified	21. NO. OF PAGES 10
	20. SECURITY CLASS (This page) Unclassified	22. PRICE A02

EPA Form 2220-1 (9-73)

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

EPA/600/D-89/110

STATISTICAL ISSUES FOR SEEDLING STUDIES IN AIR POLLUTION RESEARCH

Charles E. Peterson

Environmental Research Laboratory - Corvallis
200 S.W. 35th St.
Corvallis, OR 97333, U.S.A.

ABSTRACT

By the very nature that seedlings represent the future forests, they are an important population for exposure studies. The experimental conditions for seedling studies cover a wide range of treatment combinations, facilities, and exposure regimes. The experimental utilization of experimental studies with tree seedlings requires the approriate design and statistical concepts employed with any population response study. These issues are reviewed for their relevance to exposure studies and contribution toward studies of forest/tree response to pollutants. In order to realize the potential strengths of these studies and further enhance the intrinsic value of seedling research, the planning of experiments should include both the relevant population of interest and an explicit definition (*a priori*) of all response variables and covariates. The experimental designs should provide for pre-treatment measurements, repeated measures within growing seasons, and treatments or measurements which span more than one growing season.

INTRODUCTION

In this paper, the term "seedling" refers to trees small enough in size or stature (e.g. initial height less than 1 meter) to be used in standard open top chambers. Seedlings one to three years old have been utilized in major research programs to determine potential effects of air pollutants on important tree species (e.g. red spruce, ponderosa pine, loblolly pine, Douglas-fir, and sugar maple). Although the interest tends to focus on impacts to mature (i.e. reproductive stage) trees, information on seedlings in and of themselves, is important. Seedling studies are a means of identifying and

understanding important mechanistic processes. While a variety of experimental approaches assure a wide breadth of scientific information from seedling studies, they also pose some challenges for synthesizing and interpreting the results, both for seedling populations and extension to trees.

STATISTICAL CRITERIA

The following discussion of statistical issues is based on research funded by the U.S. Environmental Protection Agency and the USDA Forest Service. Though not unlike statistical issues found in other studies of population response, the objective of this paper is to review the following issues for their relevance to pollutant exposure studies:

1. Scope of inference. Though often overlooked, the first statistical criterion we need to identify is our population of interest, including both biologic and geographic scopes of inference. The biological level of organization in which we are interested is the seedling. Most seedling studies have provided physiological data in addition to information on relative species sensitivities. The choice of species to pick for these studies should consider selecting one set of "test" species with dissimilar physiological or ecological attributes from which to gather data for input to physiological process models, and a second set for use in validating the parameterized models.

2. Design and analysis. The primary goal of exposure research is to determine various ways in which the plant condition is affected by pollutants. It is important that the experimental design offers a reasonable chance for the objective to be fulfilled. Seedling studies are often employed to screen species or families for relative sensitivities to pollutant exposures, help build an exposure-response (E/R) model, or to help identify mechanisms of physiological response, for subsequent studies of trees. The choice of exposure levels should be documented to reflect both statistical considerations (e.g. fitting an E/R model), and biological considerations (e.g. reasonable comparison between exposure levels for mechanism of physiological response). Selection of treatments/exposure regime relevant to pollutant occurrence is always important, whether a single pollutant or pollutant co-occurrence and/or multiple stresses (Hogsett et al., 1987; Lefohn et al., 1987 and Marie and Ormrod, 1988).

(a) *Long-term studies*. Exposure of tree seedlings, as long-lived perennials, is best represented as an intervention experiment, and unlike annual crop studies, final yield as a measure of growth is not possible (Figure 1 illustrates a hypothetical response curve over time). Seedling studies rely on relative differences in growth, using repeated measures. Therefore, seedling experiments require sufficient duration and design such that treatment differences become greater than variation in annual growth. Seedlings and trees have structures and mechanisms which may either buffer response to treatments or cause responses to be delayed. Furthermore, effects of additional exposures in subsequent years needs to be addressed for trees.

(b) *Statistical power*. The probability of detecting some consequential change in condition is of major importance. Although this concern has been expressed for exposure studies (e.g. Rawlings, 1986), power is seldom reported with the experimental outcome. The number of replicates (i.e. chambers) usually influence power more so than seedlings used per treatment combination. A power curve for Pinus ponderosa is given in Figure 2. If for no other reason than repeating the experiment or planning new research, each scientist should address whether or not experimental accuracy and precision were considerations in determining sample size (i.e. the number of chambers used per treatment or the number of seedlings used per chamber).

(c) *Combining results across studies*. Statistically combining results from individual studies has the potential to increase the power of conclusions drawn from the individual experiments. Consequently, an element of regionalization can be brought into the evaluation of results. Although experimental results in forest ecology have not generally been used for regional assessments, the results from many ongoing and planned studies of forest response to acid deposition are expected to play a significant role in decision-making. Therefore, we need to formally plan experiments for this opportunity. Although the methodology may take the form of combining probabilities or test statistics (Fisher, 1932), it will more than likely involve a response surface or regression approach (e.g. Rawlings et al., 1988).

3. Response variables. All measurements or variables known to be important or of high priority, should ultimately relate to growth. Most of the seedling exposure studies address plant responses which could fall into one of three categories:

(a) *Visible Injury*. Although frequently evaluated, foliar injury has not correlated well, if at all, with growth-related changes. It is poorly defined as to scale or standard (i.e. very subjective), and quite variable (high measurement error). Therefore, while it may have limited use at a given site, it does not appear practical for use in a comprehensive analysis which might combine results across several experiments.

(b) *Growth-related changes*. Direct tree growth measurements such as stem diameter (D) and tal height (H), are commonly taken on seedlings. The repeated measures these variables affords a non-destructive means of detecting incremental changes in growth or growth patterns, in contrast to harvesting subsampl of seedlings for measuring biomass. However, the use of <u>combined</u> variables, ch as height times diameter-squared (D^2H), should be used with caution.

(c) *Physiological measurements*. Information on processes such as photosynthesis, carbon allocation, nutrient uptake, and water relations should be provided. The usefulness of seedling studies lies not only in a direct economic contribution to "risk" assessment for seedling populations in stand establishment and regeneration, but equally important is their potential to identify important physiological mechanisms which, if affected by various air pollutant scenarios, might impact tree growth.

4. <u>Covariates.</u> It is difficult to obtain plants with uniform size characteristics. With traditional ANCOVA, initial values of H and D are often employed as respective covariates for analysis of final H and final D, or incremental changes in the two variables. Before selecting a covariate, the relationship between response and the candidate covariate should be examined, to check the underlying assumptions of linearity and common slopes.

5. <u>Environmental characterization.</u> The importance of environmental characterization lies in interpreting final results. Variables selected for environmental characterization typically include temperature, relative humidity, solar radiation, and soil type, many of which cannot be simply "adjusted" for in the sense of ANCOVA.

6. Standardization of experimental methods. In addition to the issues discussed above, contrasting of results from various studies is greatly enhanced if the studies are using standard, or at least very similar, accepted methods of exposure, namely facilities and exposure regimes. Most of the current technology and research in exposure dynamics is oriented towards ozone, and based on crop studies. However, the long-lived nature of seedlings requires changes in length of exposure (post-exposure measurements and multi-year exposures), and larger chambers (for a comprehensive review on the subject, see Hogsett, et al., 1987).

7. Publishing results. The format used by investigators for reporting results commonly follows journal guidelines. Since this more abbreviated information is not always sufficient (e.g. Warren, 1986), I would suggest researchers submit (1) actual p-values (rather than *'s or NS), and (2) graphical displays of important main effects and interactions.

SUMMARY

By the very nature that seedlings represent the future forests, they are an important population in and of themselves. They are readily used in a variety of growing facilities and lend themselves to robust designs of controlled experiments. In order to realize the potential strengths of these designs, and further enhance the intrinsic value of seedling research, the planning for such experiments should:

* Specify the relevant population of interest and state to what degree the experiment will mirror that population and the environment (soil and air) in which that population exists.

* Explicitly define, *a priori*, all response variables and covariates, including how the various results will be integrated. All process or mechanistic work should be tied to growth-related changes.

* Include pre-treatment measurements, repeated measures, and treatments or measurements which span more than one growing season.

* Report results in sufficient statistical detail to allow evaluation of their individual merit and contrast with results of similar studies.

LITERATURE CITED

Fisher, R.A. 1932. Statistical Methods for Research Workers. Oliver and Boyd. Edinburgh.

Hogsett, W.E., et al. 1987. Air pollution exposure systems and experimental protocols: Volume 1: A review and evaluation of performance. EPA 600/3-87/037a; Project Officer: W.E. Hogsett, environmental research lab, Corvallis, OR

Lefohn, A.S., C.E. Davis, C.K. Jones, D.T. Tingey, and W.E. Hogsett. 1987. Co-occurrence patterns of gaseous air pollutant pairs at different minimum concentrations in the United States. Atmospheric Environment 21 (?):2435-2444.

Marie, B.A. and D.P. Ormrod. 1988. Discussion of above paper by Lefohn, et al. (1987. Atmospheric Environment 21, 2435-2444) in Atmospheric Environment 22 (6):1243.

Rawlings, J.O. 1986. Design of experiments for controlled exposure studies. In Proceedings of workshop on controlled exposure techniques and evaluation of tree responses to airborne chemicals. Technical Bulletin No. 500. National Council of the Paper Industry for Air and Stream Improvement, Inc.

Rawlings, John O., V.M. Lesser and K.A. Dassel. 1988. Statistical approaches to assessing crop losses. Environmental Pollution (In Press).

Warren, W.G. 1986. On the presentation of statistical analysis: reason or ritual. Can. Jour. For. Res. 16:1185-1191.

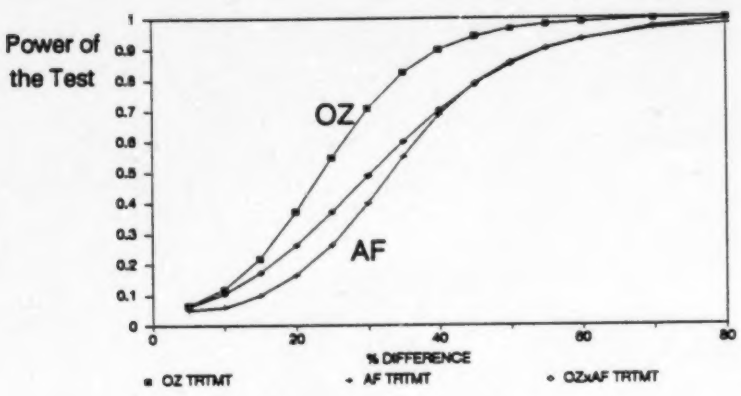

Figure 2. Power Curve constructed for ponderosa pine seedlings.

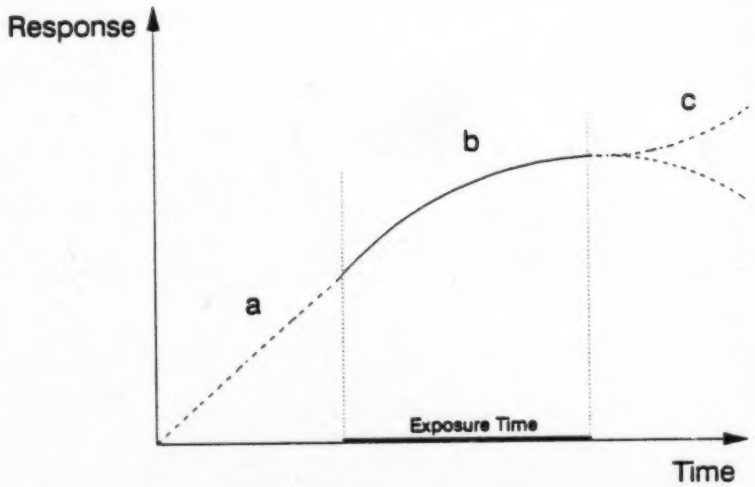

Figure 1. (a) Unobserved pre-exposure status of response variable, (b)Observed response during exposure period, (c)Possible post-exposure responses.